

A Proportional Pattern Recognition Control Scheme for Wearable A-mode Ultrasound Sensing

Xingchen Yang, Jipeng Yan, Zhenfeng Chen, Han Ding and Honghai Liu*, *Senior Member, IEEE*

Abstract—It is evident that the prevailing solution, myoelectric pattern recognition for prosthetic manipulation, constrains gesture-based interaction due to the lack of proportional control information such as exerted force. This paper reports an attempt, named simultaneous gesture recognition and muscle contraction force estimation, to realize proportional pattern recognition (PPR) control taking advantage of arm muscle deformation via wearable ultrasound sensing. We experiment with eight types of predefined hand motions, with a range of 0 - 60% maximum voluntary contraction (MVC) using a wearable multi-channel A-mode ultrasound system. The experiment result demonstrates that above 93.7% of gestures are correctly recognized during dynamic muscle contraction forces (0 - 60% MVC), albeit only training at a slight force level (<6% MVC). Besides, the adopted non-parametric Gaussian process regression estimates the muscle contraction forces accurately synchronously, with average coefficient of determination (R^2) and normalized root-mean-square error (nRMSE) of 0.927 and 0.102, respectively. These research outcomes demonstrate the feasibility of ultrasound-based PPR control, paving the way for musculature-driven applications including finer prosthetic control, remote manipulation and rehabilitation treatment.

Index Terms—Wearable ultrasound sensing, gesture recognition, force estimation, graded force control.

I. INTRODUCTION

Hand gesture recognition enables intuitive interaction with electronic and mechanical systems such as smart phone and prosthetic hand [1], [2]. Pattern recognition (PR) technique has been widely researched in this field, and desirable classification accuracy has been achieved using different sensing modalities such as surface electromyography (sEMG) [3], ultrasound [4], computer vision [5], etc. However, this classification method can only decipher discrete motion intent but cannot inherently incorporate a method of estimating the desired force or movement speed, which results in the gesture-based interaction being clumsy [6], [7].

To realize more natural gesture-based interaction, some research has focused on the inclusion of proportional control in PR control. This scheme, where proportional control information is merged into the conventional PR control for exerted force/movement speed regulation, is termed as *proportional mux* in [8], and proportional pattern recognition

(PPR) herein. In 1993, Hudgins *et al.* referred to such a control scheme preliminarily using sEMG signals, where an artificial neural network was used for gesture classification and a weighted average of the mean absolute value (MAV) of sEMG channels was used for proportional speed control [9]. Later, the superiority of this control scheme was verified by Simon *et al.* through a multi-freedom tracking task, despite the pattern classification being limited to a single degree of freedom (DOF) at a time [10]. Recently, Scheme *et al.* attempted a seven-class PPR control mission using sEMG, with result showing that the variation of muscle contraction forces during proportional control destroyed the classification performance significantly [11]. It demonstrated that sEMG was not viable in multi-class PPR control, due to the instability of myoelectric PR to variable muscle contraction intensities. Quite recently, Belyea *et al.* proposed a similar control scheme based on high density force myography, in which class-specific proportional control was realized by estimating the muscle contraction force for corresponding motions using regression technique [12]. However, the limitations of this study were that only simple hand motions were tested and the force estimation accuracy was undesirable. Consequently, merging proportional control information into conventional PR control for finer gesture interaction remains an unsolved problem.

In the last decade, ultrasound has been regarded as a promising way for muscle contraction monitoring and corresponding rehabilitation applications, owing to its ability of sensing subtle muscular deformations both in superficial and deep layers [13]–[15]. Since the muscle deformations are inherently related to hand gesture, exerted force, and movement speed, it has potential to realize desirable PPR control using ultrasound sensing. In this paper, we concentrated on the combination of dexterous gesture recognition and continuous muscle force estimation in the aim of PPR control. Early studies have reported that both finger [16], [17] and wrist motions [13], [18] can be recognized accurately using ultrasound sensing, even with robustness issues caused by arm position change, wrist rotation, and ultrasound probe shift being concerned [13], [18], [19]. However, whether ultrasound-based gesture recognition can be robust against variable muscle contraction forces remains uncertain, which is crucial for its combination with muscle contraction based proportional control. In the context of force estimation, the ability of distinguishing muscle contraction intensities using ultrasound has been reported by Ortenzi *et al.* [20]. Also, González *et al.* demonstrated that a linear relationship existed between individual finger forces and a spatial first-order feature extracted from ultrasound image [21], [22]. Yet, so far no study concentrated on the simul-

This work was supported by the National Natural Science Foundation of China (Grant Nos. 51575338, 61733011).

Yang, Yan, Chen, Ding, Liu are with the State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China. (e-mail: xingchen.yang@sjtu.edu.cn; honghai.liu@sjtu.edu.cn).

Liu is also with the School of Computing, the University of Portsmouth, PO1 3HE, Portsmouth, UK. (honghai.liu@port.ac.uk).

taneous muscle contraction force estimation during gesture recognition via ultrasound.

Notice that most of above-mentioned ultrasound-related research was based on B-mode ultrasound imaging, which was cumbersome, expensive, and unrealistic to be wearable. Our pilot study has proven that wearable A-mode ultrasound can be used for dexterous finger motion recognition, confirming the ultrasound sensing's capability for wearable applications [4]. In this paper, we utilized A-mode ultrasound to evaluate the gesture recognition performance during variable muscle contraction forces and the simultaneous muscle force estimation precision, for the purpose of wearable ultrasound based PPR control. To accurately estimate the forces of different motions, an independent model was trained for each motion instead of using a unified model, as shown in Fig. 1.

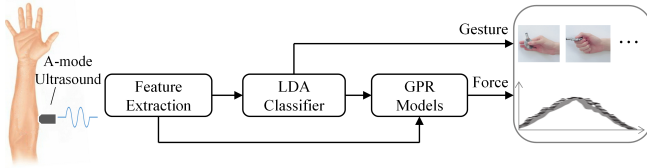


Fig. 1. Class-specific proportional control scheme. Gesture command is required to select corresponding force estimation model. LDA: Linear discriminant analysis, GPR: Gaussian process regression.

Overall, the contribution of this paper is twofold: 1) Realizing robust gesture recognition during variable muscle contraction intensities using a statistic feature; 2) Achieving accurate force estimation for given gestures using Gaussian process regression, both with wearable ultrasound modality. In conjunction with each other, proportional pattern recognition control can be realized.

The remaining parts of this paper are organized as follows: Section II-A to Section II-B provide the information of subjects and experiment protocol. Section II-C to Section II-E introduce the methods of feature extraction, pattern recognition, and feature space analysis. The force estimation models and evaluation metrics are given in Section II-F and Section II-G. Experiment results and further discussion are presented in Section III and IV.

II. METHODS

A. Subjects

Nine able-bodied subjects (all male) aged between 22 and 26 years participated in this experiment. None of them has a history of neuromuscular or joint diseases. Before participation, all subjects had received a thorough description of the experiment and signed the informed consents. The testing procedure was in accordance with the declaration of Helsinki and approved by the Ethics Committee of Shanghai Jiao Tong University.

B. Experiment Protocol

A wearable ultrasound system (US-HYS-Q2302A, Elonxi Ltd, UK) was employed in this study, which facilitated eight-channel A-mode ultrasound transducers with a frame rate of

10 Hz echo signals sequentially for each channel [23]. The transducer's driving pulse voltage was set to ± 50 V, and 1000 dots were sampled during each echo-receiving period with 20 dB amplification gain, 20 MHz sampling rate and 8 bits analog to digit conversion (ADC) resolution.

Four-channel A-mode ultrasound transducers with 5 MHz central frequency (height: 11 mm, diameter: 9 mm) were applied in this experiment. In order to secure them on the forearm, a customized holder was designed (length: 30 mm, width: 22 mm, height: 12 mm). The four-channel transducers together with the holders were equidistantly placed around the forearm with double-sided adhesive tape approximately 10 cm away from the elbow, and ultrasound gel was imposed between the skin and transducers.

Two kinds of force sensors, hand grip dynamometer and pinchmeter (G100 and P100, Biometrics Ltd., UK), were used to measure the force signals, with a sampling rate of 1000 Hz. The sensors' rated load and accuracy were 90 kg and 1% for G100, and 22.5 kg and 0.6% for P100. A software was customized for capturing ultrasound and force signals synchronously.

The experiment contained eight types of grasp motion including fine pinch (FP), middle finger pinch (MP), ring finger pinch (RF), tripod grasp (TG), four finger grasp (FG), all finger grasp (AG), lateral grasp (LG), and power grasp (PG), as shown in Fig. 2. The maximum voluntary contraction (MVC) force was tested for each motion before the experiment. In details, for each motion, the subjects were asked to exert isometric forces as big as possible for 1 minute. This procedure was repeated for three times and the average value of each motion was regarded as the MVC force. Between two MVC tests, there was a rest of 3 minutes to avoid fatigue. After the MVC test, the subjects were instructed to sit in front of a computer and grasp the force sensor, following a prompt force which linearly increased to 60% MVC in 10 seconds and then decreased to 0 in the next 10 seconds for each motion, as shown in Fig. 4. This procedure was repeated for eight trails for each motion, and there was a rest of 10 seconds between two continuous trials. The total time for each motion data collection was $(10 + 10 + 10) \times 8 = 240$ s, and $240 \times 8 = 32$ min for the whole experiment. Before the formal experiment, all the subjects were trained with eight trials or more to familiarize themselves with the experiment.

C. Signal Processing

The raw ultrasound signals were pre-processed to remove noise and enhance meaningful information, consisting of time gain compensation, band-pass filtering, envelope detection, and log compression that accorded with traditional B-mode ultrasound pre-processing [24], [25]. As mentioned above, in each frame of ultrasound echo signals, 1000 sampling dots were recorded. Since the first 20 dots and the last 20 dots carried little meaningful information, they were removed before feature extraction. The remaining 960 dots were further divided into 49 segmentations evenly, for each of which the mean value and standard deviation were calculated and noted as MSD features. The MSD features of all the 49 segments

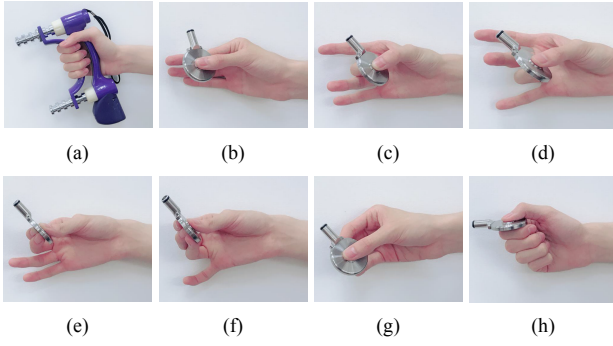


Fig. 2. Eight types of grasp motion: (a) power grasp (PG), (b) fine pinch (FP), (c) middle finger pinch (MP), (d) ring finger pinch (RF), (e) tripod grasp (TG), (f) four finger grasp (FG), (g) all finger grasp (AG), (h) lateral grasp (LG).

and four channels were concatenated into an aggregate feature vector to represent current motion state. The feature dimension was $4 \times 49 \times 2 = 392$. In order to achieve higher computing efficiency, principal component analysis (PCA) was applied to further reduce the feature dimension. The number of the selected PCA components was defined by the cumulative contribution ratio on the whole feature, which was set as 99.5%. The illustration of the MSD feature extraction is shown in Fig. 3.

In addition to the MSD feature, the spatial first-order (SFO) feature applied in [4] was also extracted for each segmentation. The dimension of the SFO feature was 392, and the dominant 99.5% of components after PCA were adopted as well.

In the case of force estimation, the force signals were firstly downsampled to 10 Hz to match the frame rate of ultrasound signals. Then, all the data were filtered using a five-order Butterworth low-pass filter with a cutoff frequency of 1 Hz. Such a filter frequency was reasonable since the muscle contraction rhythm was less than 0.05 Hz in our study.

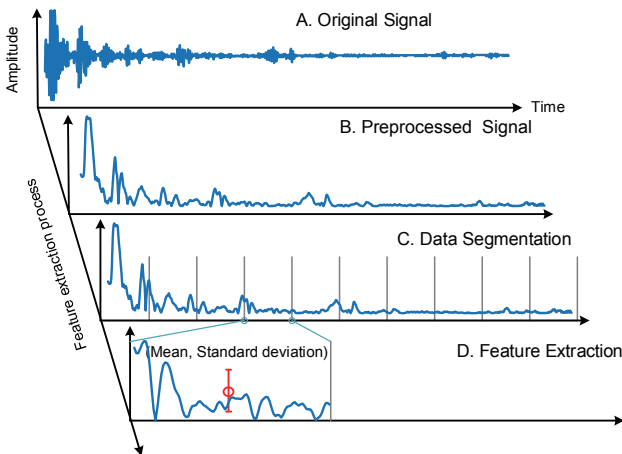


Fig. 3. Illustration of mean and standard deviation (MSD) feature extraction. Fig. 3A presents a frame of original A-mode ultrasound signal. Fig. 3B shows the signals after preprocessing. Fig. 3C shows the signals after segmentation, with a fixed segmentation length without overlap. Fig. 3D shows enlargement of a part of segmentation in Fig. 3C. The mean value and standard deviation (SD) of this part were calculated and noted as MSD feature. The MSD features of all the segmentations were concatenated together and regarded as features of the current channel.

D. Pattern Recognition

In order to evaluate the gesture recognition performance during force variation, the original dataset was divided into 10 force levels evenly, according to corresponding muscle contraction forces, as shown in Fig. 4. Each force level represented a force range of 6% MVC. During pattern recognition, the data from the force level i ($i = 1, 2, \dots, 10$) was regarded as a training set in turn, to evaluate the classification performance across all the force levels.

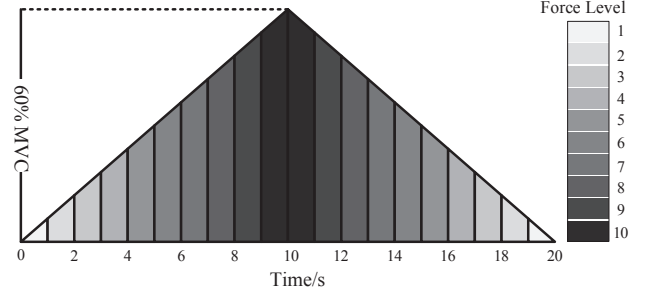


Fig. 4. Experimental paradigm and force level segmentation. For each motion, the subjects were instructed to linearly increase force to 60% maximum voluntary contraction (MVC) in 10 seconds and decrease force to 0 in the next 10 seconds. This procedure was repeated for eight trials. During pattern recognition, the recorded data was segmented into 10 force levels. Each level represented a time interval of 1 s and a force range of 6% MVC. From 1st to 10th zones, the force level increased gradually.

A linear discriminative analysis (LDA) classifier was selected for classification due to its simplicity and effectiveness [3].

E. Quantifying Feature Space

Two metrics - repeatability index (RI) [26] and safety index (SI) [21] were used to quantify feature space during force variation. The RI was formulated to measure the repeatability of feature pattern for the same motions during force variation. The SI was formulated to measure the separateness of the feature pattern for different motions during force variation. The definitions of RI and SI were given as follows. When calculating the RI or SI across two different force levels, the data from two different force levels were regarded as training data and testing data, respectively.

$$RI = \frac{1}{8} \sum_{j=1}^8 \left(\frac{1}{2} \sqrt{(u_{Trj} - u_j)^T S_{Trj}^{-1} (u_{Trj} - u_j)} \right) \quad (1)$$

where S_{Trj}^{-1} is the covariance of the training data for class j . u_{Trj} and u_j are the feature vector centroid of class j in training data and testing data, respectively. RI is one-half the Mahalanobis distance between feature vector centroid u_{Trj} and u_j , averaged across all the eight motions. A lower RI represents a greater consistency in feature pattern.

$$SI_{ij} = \frac{\max \{\sigma_{Tri}\}}{\|u_{Tri} - u_j\|}, i, j = 1, 2, \dots, 8. \quad (2)$$

where σ_{Tri} is the feature vector standard deviation of class i in training data. u_{Tri} and u_j are the feature vector centroid

of class i in training data and feature vector centroid of class j in testing data, respectively. Therefore, the SI_{ij} is the ratio between the maximum standard deviation of class i in training data over all dimensions, and the Euclidean distance between class i in training data and class j in testing data. A larger distance between u_{Tri} and u_j results in a smaller SI_{ij} .

F. Regression Models

Given a dataset \mathcal{D} of N training samples: $\mathcal{D} = \{(x_k, y_k), k = 1, 2, \dots, N\}$, where $x_k \in R^n$ is the feature vector and y_k is the corresponding force label of feature x_k . $\mathbf{x} = \{x_1, x_2, \dots, x_N\}^T$, $\mathbf{y} = \{y_1, y_2, \dots, y_N\}^T$. (x_*, y_*) is a group of testing sample.

1) *Gaussian Process Regression (GP)*: In GP, $\mathbf{y} = \{y_1, y_2, \dots, y_N\}^T$ can be imagined as a single point sampled from a multivariate (N-variate) Gaussian distribution

$$\mathbf{y} = f(\mathbf{x}) + \mathcal{N}(0, \sigma_n^2 I) \quad (3)$$

where $f(\mathbf{x}) \sim \mathcal{N}(0, K)$, $\mathcal{N}(0, \sigma_n^2 I)$ represents the random noises.

The covariance function $k(x, x') \in K$ relates to the distance between different observation values x and x' . A popular choice is to use ‘squared exponential’ kernel to represent K .

$$k(x, x') = \sigma_f^2 \exp \left[\frac{-(x - x')^2}{2l^2} \right] \quad (4)$$

where the maximum allowable covariance is defined as σ_f^2 , and l represents the width of the ‘squared exponential’ kernel function. The random noises $\mathcal{N}(0, \sigma_n^2 I)$ can be folded into $k(x, x')$, written by

$$k(x, x') = \sigma_f^2 \exp \left[\frac{-(x - x')^2}{2l^2} \right] + \sigma_n^2 \delta(x, x') \quad (5)$$

where $\delta(x, x')$ is the Kronecker delta function.

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_N) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_N, x_1) & k(x_N, x_2) & \dots & k(x_N, x_N) \end{bmatrix} \quad (6)$$

Given testing data (x_*, y_*) , all the data can be represented as a sample from a new multivariate Gaussian distribution

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} K & K^T \\ K_* & K_{**} \end{bmatrix} \right) \quad (7)$$

$$K_* = [k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_N)], K_{**} = k(x_*, x_*). \quad (8)$$

The probability $p(y_*|\mathbf{y})$ also follows a Gaussian distribution

$$y_*|\mathbf{y} \sim \mathcal{N}(K_* K^{-1} \mathbf{y}, K_{**} - K_* K^{-1} K_*^T). \quad (9)$$

The best estimation for y_* is the mean value of this distribution, and the uncertainty of the estimation is its variance

$$\tilde{y}_* = K_* K^{-1} \mathbf{y} \quad (10)$$

$$\text{var}(y_*) = K_{**} - K_* K^{-1} K_*^T \quad (11)$$

2) *Linear Regression (LR)*: Classical LR was also evaluated in this context, where the estimation value of y_* is given by

$$\tilde{y}_* = x_*^T (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} \quad (12)$$

G. Force Estimation Metrics

The above-mentioned two regression models were applied to predict the muscle contraction forces. For each motion, a regression model was trained with data from the first six trials and tested with data from the last two trials, which intended to simulate the real-time application scenario.

The coefficient of determination (R^2) and normalized root-mean-square error ($nRMSE$) were applied to quantify the force estimation precision, which were defined as

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

$$nRMSE = \frac{\sqrt{\sum_{i=1}^n (y_i - \tilde{y}_i)^2 / n}}{y_{\max} - y_{\min}} \quad (14)$$

where n is the number of test samples, y_i is the actual force value of the i th sample, \tilde{y}_i is the corresponding estimated force. \bar{y} , y_{\max} , y_{\min} are the average force value, maximum force value, and minimum force value over all the testing samples.

III. EXPERIMENT RESULTS

A. Gesture Recognition Accuracy

Fig. 5 shows the gesture recognition accuracy of the two types of feature during variable force levels. The result was averaged across different subjects and motions. From F1 to F10, the classifiers were trained with data from force level 1 to 10, and tested across all the force levels (0 - 60% MVC). It was obvious that the proposed MSD feature was more robust against force variation compared with the SFO feature, especially when training at lower force levels, such as F1, F2. Paired t test revealed that the difference between the MSD feature and the SFO feature was significant, when training at F1, F2, F6, and F9. The best classification accuracy can be achieved for both the MSD feature and the SFO feature when training with data from F3 (12% - 18% MVC), with average classification accuracy of 98.7% and 97.7%, respectively. Moreover, for the MSD feature, regardless of the training force level, a remarkable classification accuracy (>93.7%) can be achieved during variable force levels.

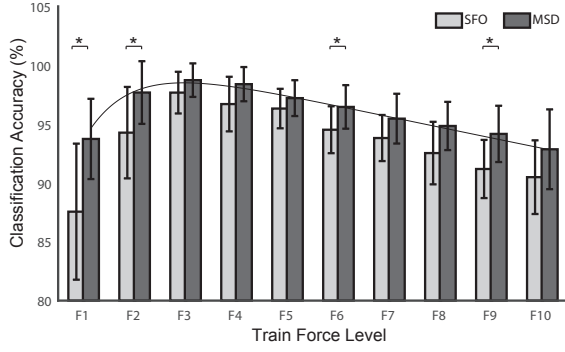


Fig. 5. Classification accuracy of different features during variable force levels. From F1 to F10, the classifiers are trained with data from force level 1 to 10, and tested across all the force levels. The result is averaged across different subjects and motions, and standard deviations are shown with error bars. Statistical analysis results for different features are given above the histogram, where sign * denotes $p < 0.05$.

B. Feature Space Analysis during Force Variation

Fig. 6 presents the average feature repeated index (RI) between force level i ($i = 1, 2, \dots, 9$) and force level 10 (maximum force level) across different subjects. The MSD feature was analyzed here considering its superior performance. The data from the maximum force level and force level i was regarded as training data and testing data respectively, when calculating the RI defined in Eq. (1). As the force level approached to maximum force level, the RI decreased gradually. It revealed that the MSD features varied with forces, since the RI is a metric of consistency in the feature pattern. Notice that the RI of F1 was lower than that of F2. A potential reason might be that it was hard to control the exerted force during initial muscle contraction, which caused inevitable force fluctuation in F1. Thus, the feature similarity between F1 and F10 was a little higher than that of F2 and F10.

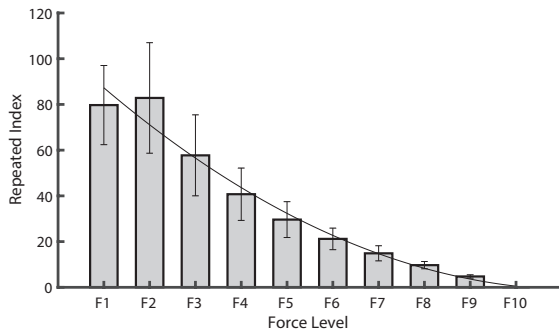


Fig. 6. Average repeated index (RI) between force level i ($i = 1, 2, \dots, 9$) and force level 10 for the MSD feature across different subjects. SDs are shown with error bars. A lower RI indicates a smaller intraclass distance.

Fig. 7 shows the average safe index (SI) between maximum force level and minimum force level for the MSD feature across different subjects. Correspondingly, the data from maximum force level and minimum force level were regarded as training data and testing data respectively, for the calculation of the SI defined in Eq. (2). For each row, the principal diagonal element was larger than non-diagonal elements. It

demonstrated that the intraclass feature distances were smaller than interclass feature distances even under the circumstance of the largest force variation, since a larger class distance resulted in a smaller SI. By comprehensive analysis of RI and SI, it was speculated that the MSD features varied with muscle contraction forces, while intraclass feature distances were always smaller than interclass feature distances, which guaranteed the robust gesture recognition performance during force variation. On the other hand, now that the feature shifted with force variation, it might be feasible to estimate forces using ultrasound features synchronously.

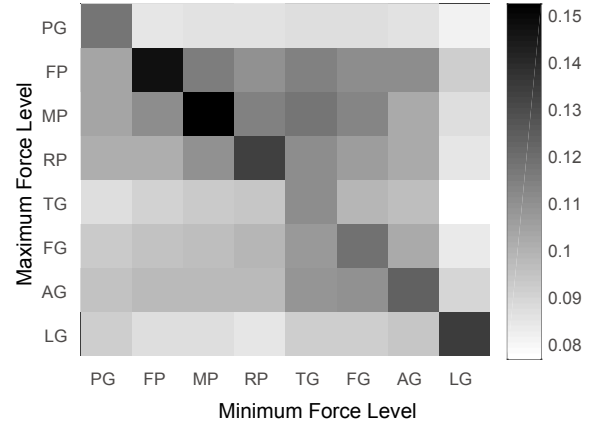


Fig. 7. Average safe index (SI) between maximum force level (F10) and minimum force level (F1) for the MSD feature across different subjects. A smaller SI value indicates a larger feature distance.

C. Force Estimation Performance

Fig. 8 compares the force estimation performance between different combinations of feature and regression method. According to Fig. 8A, the average R^2 coefficients between actual and estimated forces were 0.851 ± 0.075 for SFO feature with LR model, 0.899 ± 0.054 for SFO feature with GP model, 0.905 ± 0.057 for MSD feature with LR model, and 0.927 ± 0.045 for MSD feature with GP model, respectively. Overall, the MSD feature combined with GP model outperformed the other combinations. A paired t test revealed that the force estimation precision of the MSD feature was significantly better than the SFO feature, no matter using LR or GP models ($p < 0.001$). Also, the GP outperformed the LR significantly for either MSD or SFO features ($p < 0.05$).

According to Fig. 8B, the average $nRMSE$ coefficients between actual and estimated grasp forces were 0.141 ± 0.040 for SFO feature with LR model, 0.124 ± 0.043 for SFO feature with GP model, 0.112 ± 0.037 for MSD feature with LR model, and 0.102 ± 0.037 for MSD feature with GP model, respectively. Likewise, the combination of the MSD feature and GP model performed best over all the combinations. A paired t test showed that the performance of the SFO feature and the MSD feature was significantly different, for either LR or GP models ($p < 0.003$). Also, the GP outperformed LR significantly, when using the SFO feature ($p < 0.05$). However,

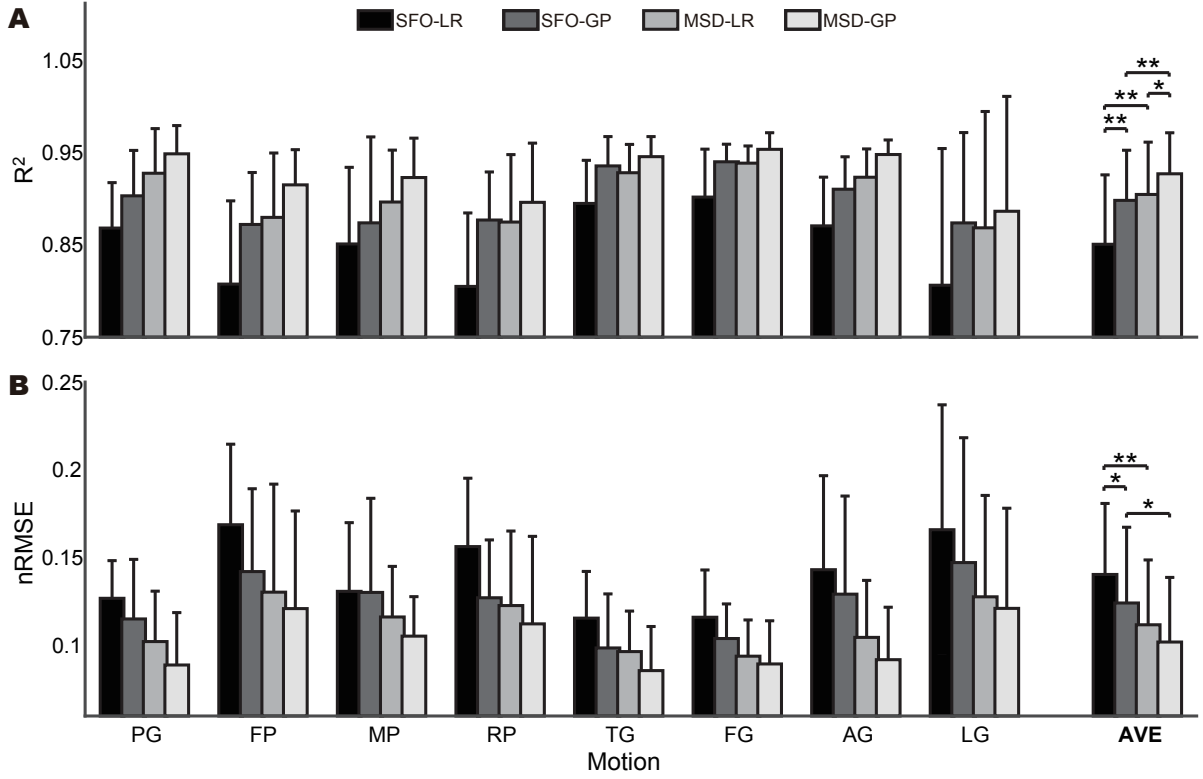


Fig. 8. The R^2 and $nRMSE$ coefficients of class-specific force estimation, where data is averaged across all the subjects. The last column (AVE) denotes the average values across different motions and subjects. The Gaussian process regression and linear regression are denoted as GP and LR, respectively. Standard deviation is represented with error bar. Statistical analysis results for different features and regression models are shown above the last column, where sign * denotes $p < 0.05$ and sign ** denotes $p < 0.001$.

the performance of LR was comparable to GP when using the MSD feature ($p = 0.14$).

Fig. 9 shows a representative example of force estimation performance when using the MSD feature and GP model. It was evident that the estimated forces were in good agreement with the actual forces for all the eight grasp motions. However, a limitation was that the estimated error increased as the force magnitude increased, for the motions like FP, MP, and FG.

IV. CONCLUDING REMARKS

This paper demonstrated that simultaneous gesture recognition and muscle contraction force estimation can be achieved using wearable A-mode ultrasound sensing, on the basis of a statistical feature and a non-parametric Gaussian process regression model. The average gesture recognition accuracy across variable force levels was above 93.7%, albeit only training at a slight force level ($<6\%$ MVC). Moreover, when training with forces ranged from 12% to 18% MVC, the average classification accuracy can be improved to 98%. More importantly, the exerted forces for corresponding motions can be estimated accurately synchronously, with average R^2 and $nRMSE$ of 0.927 and 0.102, respectively. This study opens up a scenario in which users can control a prosthetic hand or smart phone in a much finer way, since not only the intended motion but also the amount of force can be outputted accurately.

A. Method Considerations

In the part of gesture recognition, the performance of the SFO feature and the MSD feature was compared, in terms of the classification accuracy across variable force levels. While satisfying performance has been achieved by the SFO feature [4], we found that the MSD feature can further improve the classification performance in the presence of variable muscle contraction forces. As shown in Fig. 5, remarkable classification accuracy ($>93.7\%$) can be achieved by the MSD feature, when only training at a slight force level ($<6\%$ MVC) and testing across a large range of force levels (0 - 60% MVC), which was significantly better than the SFO feature. As training with forces ranged from 12% - 18% MVC, the best classification accuracy can be achieved for both the MSD feature and the SFO feature. Further increasing the training force would spoil the classification performance. This result provides a recommendation to user training that exerting a mild force is adequate for the training, by which the training burden and risk of fatigue might be reduced.

To further validate the robustness of the MSD feature against force variation, we segmented the collected dataset into 5 and 20 force levels respectively, and only one of them was used for training (Fig. 10). In the case of 5 force levels, the classification accuracy was higher than 94.5% regardless of the training force level. In the case of 20 force levels, the classification accuracy was higher than 90.7% albeit only training at 0 - 3% MVC. In both cases, the training force

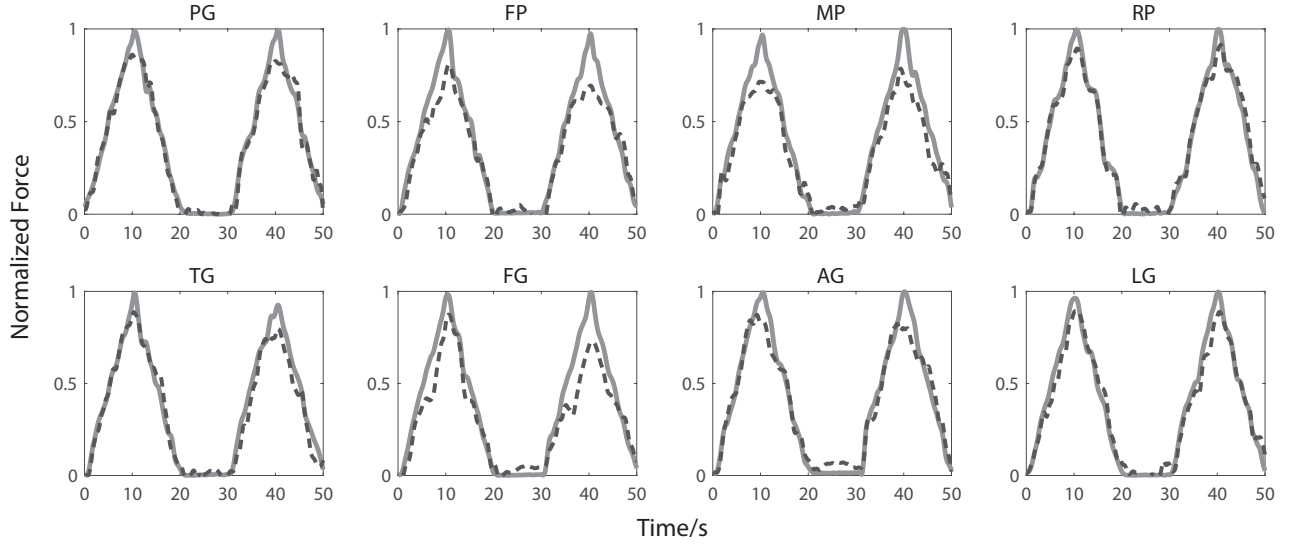


Fig. 9. Force estimation results for different motions using MSD feature and GP model. The forces were normalized to 0 - 1 for each motion.

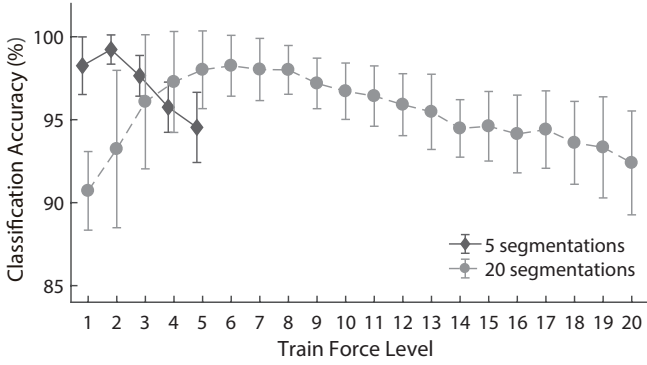


Fig. 10. Classification accuracy when segmenting the dataset into 5 and 20 force levels. In both cases, a LDA classifier was trained using data from each force level and tested across all the force levels. The horizontal axis denotes the training force levels, and the vertical axis denotes the test accuracy. Result is averaged across different subjects and motions, and standard deviations are shown with error bars.

level with the highest classification accuracy was around 18% MVC.

According to feature space analysis, it was found that the MSD feature shifted with force variation (Fig. 6), while the intraclass distances were always smaller than interclass distances (Fig. 7). This explained why the MSD feature was robust against force variations.

In the part of force estimation, the MSD feature performed better as well whether using LR or GP models (Fig. 8). Besides, it was found that the MSD feature was approximately linear to muscle contraction force, since a linear regression model can predict the exerted force precisely by the MSD feature ($R^2 = 0.905$, $nRMSE = 0.112$). The linear relationship also existed between the SFO feature and the muscle contraction force, with a relatively lower degree of linearity ($R^2 = 0.899$, $nRMSE = 0.124$). This may be attributed to the fact that the muscle deformation was approximately linear to the contraction force [27], [28], and ultrasound signals can reflect the muscle deformations with high precision. Apart

from this, we found GP regression can improve the force estimation accuracy significantly. The potential reason was that the relationship between the muscle deformation and the generated force was not absolutely linear, and the advantage of GP in non-linear regression has been demonstrated in previous research [29]. Only constrained muscle force estimation (i.e. following a triangle prompt force) was evaluated in this work. However, the ability of estimating force via ultrasound under random muscle contraction has been demonstrated, albeit only training with zero and maximum muscle contraction forces [21].

The latency of data processing during prediction was summarized in Table I. The overall consumed time was 7.935 ms for the MSD feature and 8.635 ms for the SFO feature, indicating that the system can be applied in real-time sceneries where subjects cannot perceive the system delay [3].

The training consumption of current research was a little high (>32 min). But according to the experiment result, the training burden for the gesture recognition can be reduced, since only training at a slight force level was enough. Additionally, the training efficiency for the force estimation can be improved, because simple on-off training strategy has been proven to be effective [21].

TABLE I
LATENCY OF DATA PROCESSING (MS)

	Feature Ext	PCA	LDA	GP	All
SFO	5.6	0.015	0.02	3	8.635
MSD	4.2	0.015	0.02	3.7	7.935

Feature Ext: Feature Extraction

B. Comparison with Related Work

To our knowledge, the concept of simultaneous gesture recognition and force estimation for PPR control was origi-

nally proposed by Hudgins *et al.*, who tried to classify four types of limb motions using transient sEMG signals (about 200 ms) and estimate the muscle contraction intensity using MAV of sEMG signals subsequently [9]. However, the classification performance of this scheme was not desirable in this paper, since only 200 ms transient myoelectric signals were applied for classification. Later, Castellini *et al.* conducted a similar research via sEMG, to classify four grasp motions using pattern recognition approach and predict corresponding forces synchronously using regression technique [30], [31]. Results showed that reasonable accuracy can be attained, with average gesture recognition accuracy and force estimation error of 89.67% and 7.89%, respectively. In their subsequent work, they found that comparable results can be achieved by three amputees as well [32]. However, only four/five grasp motions were evaluated in these research and training with variable forces for each motion was required for the sake of desirable classification accuracy, which would increase the training and computing burden.

Scheme *et al.* also reported a PPR control experiment using sEMG [11]. Typically, the intensity of the sEMG signals was used to dictate the exerted force for a target tracking task, and a LDA classifier was applied for classifying seven types of tracking motion including rest. It was demonstrated that the variation of muscle contraction forces during proportional control influenced the classification accuracy dramatically. In the case of training and testing with data from different force levels, the average classification accuracy was only 31.2%. With a strategy of training with forces ramped from 20% to 80% MVC, the classification error can be decreased to 18.9%. The deficiency of study was that the gesture recognition accuracy was undesirable during proportional control, despite training with data from various force levels.

Additionally, Belyea *et al.* attempted a class-specific proportional control scheme using high density force myography, where a support vector machine classifier was used to classify four/six hand motions and support vector regression was applied to estimate the exerted force for given motions [12]. While remarkable classification accuracy (99%) was achieved, only simple hand motions (i.e. wrist motions and hand open/close) were studied in this paper and the force estimation precision was not satisfying ($R^2 = 0.83$).

Compared with the above-mentioned research, our ultrasound-based method can achieve more robust gesture recognition performance during force variation and comparable force estimation precision than sEMG [31], [33], with more dexterous gestures being considered. It is supposed to be a superior method for the intuitive PPR control. The key advantage of the ultrasound-based sensing method is that both the muscle deformation during motion switching [16] and force variation [27] can be captured clearly by ultrasound sensing. This information can be well represented by the MSD feature proposed in this paper. As shown in feature space analysis, the MSD feature changed with force variation and motion switching (Fig. 6, Fig. 7). Moreover, the change during force variation was slighter than motion switching (Fig. 7), which guaranteed that the intraclass distances were always smaller than interclass

distances. Correspondingly, robust gesture recognition and accurate force estimation can be realized synchronously via ultrasound.

C. Potential Applications

The application of the results shown in this paper is potentially multifarious, in the areas of prosthetic hand control, remote manipulation, human motion analysis, neuropathy rehabilitation, etc. In contrast to a conventional pattern recognition system, our proposed approach can realize a much finer control, since not only discrete gestures but also continuous forces can be given. It is of vital importance for some real-life applications. Take a certain case of grabbing an egg with a prosthetic hand as an example, the control system should be able to enforce correct grasp motions as well as a reasonable amount of force to lift the egg without breaking it [31]. It is feasible to solve this problem with our proposed method. However, conventional pattern recognition cannot handle it due to the lack of force information. Additionally, still in the region of prosthetic hand control, the continuous force information can be applied for speed control, making the movement of the prosthetic hand more fluent and natural.

Finally, only able-bodied subjects participated in this experiment. The viability of this approach for transradial amputees would be further analyzed for its clinical applications.

REFERENCES

- [1] S. Jiang, B. Lv, W. Guo, C. Zhang, H. Wang, X. Sheng, and P. Shull, "Feasibility of wrist-worn, real-time hand and surface gesture recognition via sEMG and imu sensing," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3376–3385, 2018.
- [2] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *Journal of Rehabilitation Research and Development*, vol. 48, no. 6, p. 643, 2011.
- [3] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848–854, 2003.
- [4] X. Yang, X. Sun, D. Zhou, Y. F. Li, and H. Liu, "Towards wearable a-mode ultrasound sensing for real-time finger motion recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 6, pp. 1199–1208, 2018.
- [5] N. H. Dardas and N. D. Georganas, "Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 11, pp. 3592–3607, 2011.
- [6] E. Scheme, B. Lock, L. Hargrove, W. Hill, U. Kuruganti, and K. Englehart, "Motion normalized proportional control for improved pattern recognition-based myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 149–157, 2014.
- [7] N. Jiang, S. Dosen, K.-R. Muller, and D. Farina, "Myoelectric control of artificial limbs: is there a need to change focus?" *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 152–150, 2012.
- [8] A. Fougner, O. Stavdahl, P. J. Kyberd, Y. G. Losier, and P. A. Parker, "Control of upper limb prostheses: terminology and proportional myoelectric control—a review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 663–677, 2012.
- [9] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.
- [10] A. M. Simon, K. Stern, and L. J. Hargrove, "A comparison of proportional control methods for pattern recognition control," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 3354–3357.

- [11] E. Scheme and K. Englehart, "Training strategies for mitigating the effect of proportional control on classification in pattern recognition based myoelectric control," *J Prosthet Orthot*, vol. 25, no. 25, pp. 76–83, 2013.
- [12] A. T. Belyea, K. B. Englehart, and E. J. Scheme, "A proportional control scheme for high density force myography," *Journal of Neural Engineering*, vol. 15, no. 4, 2018.
- [13] N. Akhlaghi, C. A. Baker, M. Lahlou, H. Zafar, K. G. Murthy, H. S. Rangwala, J. Kosecka, W. M. Joiner, J. J. Pancrazio, and S. Sikdar, "Real-time classification of hand motions using ultrasound imaging of forearm muscles," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1687–1698, 2016.
- [14] Y. Huang, X. Yang, Y. Li, D. Zhou, K. He, and H. Liu, "Ultrasound-based sensing models for finger motion classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1395–1405, 2018.
- [15] X. Sun, X. Yang, X. Zhu, H. Liu, and H. Liu, "Dual-frequency ultrasound transducers for the detection of morphological changes of deep-layered muscles," *IEEE Sensors Journal*, vol. 18, no. 4, pp. 1373–1383, 2018.
- [16] J. Shi, S.-x. Hu, Z. Liu, J.-Y. Guo, Y.-j. Zhou, and Y.-p. Zheng, "Recognition of finger flexion from ultrasound image with optical flow: A preliminary study," in *Biomedical Engineering and Computer Science (ICBECS), 2010 International Conference on*. IEEE, 2010, pp. 1–4.
- [17] S. Sikdar, H. Rangwala, E. B. Eastlake, I. A. Hunt, A. J. Nelson, J. Devanathan, A. Shin, and J. J. Pancrazio, "Novel method for predicting dexterous individual finger movements by imaging muscle activity using a wearable ultrasonic system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 69–76, 2014.
- [18] J. McIntosh, A. Marzo, M. Fraser, and C. Phillips, "EchoFlex: Hand gesture recognition using ultrasound imaging," in *CHI Conference on Human Factors in Computing Systems*, 2017, pp. 1923–1934.
- [19] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 788–797, 2012.
- [20] V. Ortenzi, S. Tarantino, C. Castellini, and C. Cipriani, "Ultrasound imaging for hand prosthesis control: a comparative study of features and classification methods," in *IEEE International Conference on Rehabilitation Robotics*, 2015, pp. 1–6.
- [21] D. S. González and C. Castellini, "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees," *Frontiers in neurobotics*, vol. 7, 2013.
- [22] C. Castellini and D. S. González, "Ultrasound imaging as a human-machine interface in a realistic scenario," vol. 40, no. 6, pp. 1486–1492, 2013.
- [23] N. Hettiarachchi, Z. Ju, and H. Liu, "A new wearable ultrasound muscle activity sensing system for dexterous prosthetic control," in *IEEE International Conference on Systems, Man, and Cybernetics*, 2016, pp. 1415–1420.
- [24] G. Farhat, *Diagnostic Ultrasound Imaging: Inside Out*. Lippincott-Raven, 2007.
- [25] Y. Li, K. He, X. Sun, and H. Liu, "Human-machine interface based on multi-channel single-element ultrasound transducers: A preliminary study," in *IEEE International Conference on E-Health Networking, Applications and Services*, 2016, pp. 1–6.
- [26] N. E. Bunderson and T. A. Kuiken, "Quantification of feature space changes with experience during electromyogram pattern recognition control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, p. 239, 2012.
- [27] S. Jun, Z. Yong-Ping, H. Qing-Hua, and C. Xin, "Continuous monitoring of sonomyography, electromyography and torque generated by normal upper arm muscles during isometric contraction: sonomyography assessment for arm muscles," *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 1191–1198, 2008.
- [28] A. Takashi, B. R. Counts, B. E. Barnett, S. J. Dankel, L. Kofan, and J. P. Loenneke, "Associations between handgrip strength and ultrasound-measured muscle thickness of the hand and forearm in young men and women," *Ultrasound in Medicine and Biology*, vol. 41, no. 8, pp. 2125–2130, 2015.
- [29] C. E. Rasmussen, *Evaluation of gaussian processes and other methods for non-linear regression*. University of Toronto, 1997.
- [30] C. Castellini, P. V. D. Smagt, G. Sandini, and G. Hirzinger, "Surface emg for force control of mechanical hands," in *IEEE International Conference on Robotics and Automation*, 2008, pp. 725–730.
- [31] C. Castellini and P. V. D. Smagt, *Surface EMG in advanced hand prosthetics*. Springer-Verlag New York, Inc., 2009.
- [32] C. Castellini, E. Gruppioni, A. Davalli, and G. Sandini, "Fine detection of grasp force and posture by amputees via surface electromyography," *Journal of Physiology - Paris*, vol. 103, no. 3, pp. 255–262, 2009.
- [33] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use," *Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 643–659, 2011.